

QCD Jet Rates with the Inclusive Generalized k_t Algorithms

Erik Gerwick, Steffen Schumann

II. Physikalisches Institut, Universität Göttingen, Germany

Ben Gripaios, Bryan Webber

Cavendish Laboratory, University of Cambridge, U.K.

ABSTRACT: We derive generating functions, valid to next-to-double logarithmic accuracy, for QCD jet rates according to the inclusive forms of the k_t , Cambridge/Aachen and anti- k_t algorithms, which are equivalent at this level of accuracy. We compare the analytical results with jet rates and average jet multiplicities from the SHERPA event generator, and study the transition between Poisson-like and staircase-like behaviour of jet ratios.

KEYWORDS: [qcd](#), [jet](#).

Contents

1. Introduction	2
2. The inclusive generalized k_t jet algorithms	3
3. Sudakov factors	4
4. Generating functions	5
4.1 Next to double leading logarithms	5
4.2 Leading double logarithms	6
5. Jet rates	6
6. Average jet multiplicity	10
6.1 Next to double leading logarithms	10
6.2 Leading double logarithms	11
6.3 Running coupling	12
6.4 Numerical solution: DLA	12
6.5 Numerical solution: NDLA	13
7. Monte Carlo results: e^+e^-	13
8. Monte Carlo results: pp	16
9. Jet scaling patterns	17
9.1 Poisson breaking components	18
9.2 R -dependence of scaling	19
10. Conclusions	20
A. Derivation of the PDE for the average jet multiplicity	21
B. Properties of the PDE for the average jet multiplicity	22

1. Introduction

The production of hadronic jets in high-energy collisions is one of the most striking and universal features of particle physics. It is now well understood that jets are associated with hard parton emission, followed by parton showering and hadronization. However, the observed features of jets depend not only on their intrinsic properties but also on the algorithms used to find them. These algorithms have been the object of many years of development and refinement – see [1] for a review.

The jet algorithms used most widely nowadays are of the sequential recombination type. Of these, the k_t algorithm [2–4] was used extensively at LEP and HERA, and to a limited extent at the Tevatron. The good theoretical properties of the k_t algorithm mean that jet rates and multiplicities can be calculated perturbatively, either in fixed order or in an all-orders logarithmic approximation. However, the irregular angular shapes of k_t -jets make them less suitable for analyses of hadron collider events, where multiple parton interactions give rise to underlying activity not associated with the primary hard process.

A variant of the k_t algorithm that allows better control of the underlying event is anti- k_t [5], which has become the preferred jet finder at the LHC. It belongs to the general category of inclusive generalized k_t jet algorithms, defined in [6]. Many of the good theoretical features of k_t extend more broadly to the members of this category. In particular their leading and next-to-leading double logarithmic jet rates can be calculated to all orders, and are in fact the same for all members, including k_t and anti- k_t .

In the present paper we calculate these jet rates, along with the average jet multiplicities, in a variety of approximations. We do so by finding, in each case, the equations that the quark and gluon jet generating functions [7–9] must satisfy. In some cases, an explicit solution in terms of special functions can be found, while in others we are forced to resort to numerical methods (though we can get some analytic understanding of the properties of the solution, as we describe in detail for one case in an Appendix). We then compare the results of our theoretical calculations with those obtained via Monte Carlo simulations using the SHERPA event generator [10, 11]. Our theoretical calculations of jet rates are carried out for the case of e^+e^- collisions, but we also perform calculations and simulations of the average jet multiplicity in hadron-hadron collisions, finding that similar behaviour is obtained.

The outline is as follows. We begin by recalling the definitions of the inclusive jet algorithms and the relevant Sudakov form factors. In Section 4, we introduce the jet generating functions and derive the integral equations that they satisfy at double-leading-logarithmic accuracy (DLA) and next-to-double-leading-logarithmic accuracy (NDLA). In Sections 5 and 6, we compute the resulting jet rates and average jet multiplicities, both at fixed coupling and also at next-to-leading order in the running coupling. In Sections 7 and 8, we compare with results of simulations

performed using SHERPA for e^+e^- and pp collisions, respectively. In Section 9 we consider the implications for the scaling patterns of jet multiplicities. Our conclusions are summarized in Section 10. Details of the derivation and properties of the partial differential equation (PDE) for the average jet multiplicity are relegated to appendices.

2. The inclusive generalized k_t jet algorithms

We consider first the case of multijet production in e^+e^- annihilation, for which the inclusive algorithms are defined as described in the FastJet user manual [6], Sect. 4.5. The distance measures are

$$\begin{aligned} d_{ij} &= \min\{E_i^{2p}, E_j^{2p}\} \frac{(1 - \cos \theta_{ij})}{(1 - \cos R)} \\ &\equiv \min\{E_i^{2p}, E_j^{2p}\} \xi_{ij} / \xi_R, \\ d_{iB} &= E_i^{2p}, \end{aligned} \tag{2.1}$$

with $p = 1, 0, -1$ for the k_t [2], Cambridge/Aachen [12, 13] and anti- k_t [5] algorithms, respectively. At any stage of clustering, if a d_{ij} is the smallest measure we combine objects i and j . If d_{iB} is the smallest we call i a jet candidate and remove it from the clustering list. We then call jet candidates with energy $E_i > E_R$ resolved jets. Thus the jet rates, at a given value of the e^+e^- centre-of-mass energy E_{cm} , are functions of the radius parameter R and the minimum jet energy E_R . This is in contrast to the *exclusive* k_t (Durham) algorithm, where one effectively sets $\xi_R = \frac{1}{2}$ and continues clustering objects until all d_{ij} are above some fixed value $d_{\text{cut}} = y_{\text{cut}} E_{\text{cm}}^2$, so that the jet rates are functions of the single parameter y_{cut} .

In hadron-hadron collisions, the c.m. frame of the parton-parton hard scattering process is not known and therefore one has to adopt a longitudinally invariant form of the algorithms [3, 4]. To that end, Eqs. (2.1) are replaced by

$$\begin{aligned} d_{ij} &= \min\{p_{ti}^{2p}, p_{tj}^{2p}\} \frac{\Delta R_{ij}^2}{R^2}, \\ \Delta R_{ij}^2 &\equiv (y_i - y_j)^2 + (\phi_i - \phi_j)^2, \\ d_{iB} &= p_{ti}^{2p}, \end{aligned} \tag{2.2}$$

where p_{ti} , y_i and ϕ_i are the transverse momentum, rapidity and azimuth of object i , respectively, and we define jet candidates with $p_{ti} > E_R$ as resolved jets.

As far as leading logarithms are concerned, the jet rates defined by (2.1) and (2.2) will be the same, and therefore in the following Sections we refer mainly to e^+e^- annihilation. By “leading logarithms” here we always mean leading double and next-to-double logarithms, $\alpha_s^n \log^{2n}$ and $\alpha_s^n \log^{2n-1}$, where the logarithms are those of $1/R$ and/or Q/E_R , Q being the hard process scale. By taking E_R sufficiently large

in hadron-hadron collisions, we avoid such leading contributions from initial-state showering and the underlying event, so these terms are determined by the timelike showering of final-state partons.

3. Sudakov factors

The evolution scale for coherent parton showering is $\xi \equiv 1 - \cos \theta$ with θ the emission angle. To be resolved, an emission must have $\xi > \xi_R$ and $E > E_R$. The probability for a single resolvable gluon emission from a quark of energy E at scale ξ is thus

$$\mathcal{P}_q(E, \xi) = \int_{\xi_R}^{\xi} \frac{d\xi'}{\xi'} \int_{E_R/E}^1 dz \frac{\alpha_S(k_t^2)}{2\pi} P_{gq}(z) , \quad (3.1)$$

where the running coupling is evaluated at the transverse momentum scale of the emission, $k_t^2 = z^2 E^2 \xi'$,

$$\frac{\alpha_S(k_t^2)}{\pi} = \frac{1}{b_0 \ln(z^2 E^2 \xi' / \Lambda^2)} \quad (3.2)$$

with $b_0 = (11C_A - 2n_f)/12$. Defining $\bar{\alpha}_S = \alpha_S(E^2 \xi)/\pi$, i.e. in terms of the coupling at the hard scale, we have to next-to-double-log accuracy (NDLA)

$$\frac{\alpha_S(k_t^2)}{\pi} = \bar{\alpha}_S - b_0 \bar{\alpha}_S^2 \left[2 \ln z + \ln \left(\frac{\xi'}{\xi} \right) \right] \quad (3.3)$$

and

$$\begin{aligned} \mathcal{P}_q(E, \xi) = & C_F \bar{\alpha}_S \ln \left(\frac{\xi}{\xi_R} \right) \left[\ln \left(\frac{E}{E_R} \right) - \frac{3}{4} \right] + \\ & \frac{1}{2} C_F b_0 \bar{\alpha}_S^2 \ln \left(\frac{\xi}{\xi_R} \right) \ln \left(\frac{E}{E_R} \right) \left[2 \ln \left(\frac{E}{E_R} \right) + \ln \left(\frac{\xi}{\xi_R} \right) \right] . \end{aligned} \quad (3.4)$$

Then the probability for no resolvable emissions (the quark Sudakov factor) is

$$\Delta_q(E, \xi) = \exp [-\mathcal{P}_q(E, \xi)] . \quad (3.5)$$

Similarly for a gluon, the probability of a single resolvable gluon, quark or antiquark emission is

$$\mathcal{P}_g(E, \xi) = \int_{\xi_R}^{\xi} \frac{d\xi'}{\xi'} \int_{E_R/E}^1 dz \frac{\alpha_S(k_t^2)}{2\pi} [P_{gg}(z) + P_{qg}(z)] , \quad (3.6)$$

which gives to NDLA

$$\begin{aligned} \mathcal{P}_g(E, \xi) = & \bar{\alpha}_S \ln \left(\frac{\xi}{\xi_R} \right) \left[C_A \ln \left(\frac{E}{E_R} \right) - b_0 \right] + \\ & \frac{1}{2} C_A b_0 \bar{\alpha}_S^2 \ln \left(\frac{\xi}{\xi_R} \right) \ln \left(\frac{E}{E_R} \right) \left[2 \ln \left(\frac{E}{E_R} \right) + \ln \left(\frac{\xi}{\xi_R} \right) \right] , \end{aligned} \quad (3.7)$$

and the gluon Sudakov factor is

$$\Delta_g(E, \xi) = \exp [-\mathcal{P}_g(E, \xi)] . \quad (3.8)$$

Note that all this is independent of the value of p , so that all the inclusive generalized k_t algorithms are equivalent at this level of precision.

4. Generating functions

By definition the generating function for resolved jets from a quark ($i = q$) or gluon ($i = g$) of energy E at scale ξ is [7–9]

$$\Phi_i(u, E, \xi) = \sum_{n=0}^{\infty} u^n R_n^i(E, \xi) , \quad (4.1)$$

where R_n^i is the corresponding n -jet rate, i.e. the probability of finding n resolved jets. Thus the jet rates can be recovered from the generating function by successive differentiation at $u = 0$:

$$R_n^i(E, \xi) = \frac{1}{n!} \frac{\partial^n}{\partial u^n} \Phi_i(u, E, \xi)|_{u=0} . \quad (4.2)$$

On the other hand the average multiplicity of resolved jets is obtained by differentiating at $u = 1$. Writing the average jet multiplicity from a quark or gluon of energy E at scale ξ as $\mathcal{N}_i(E, \xi)$, we have

$$\mathcal{N}_i(E, \xi) = \sum_{n=0}^{\infty} n R_n^i(E, \xi) = \frac{\partial}{\partial u} \Phi_i(u, E, \xi)|_{u=1} . \quad (4.3)$$

The generating functions $\Phi_{q,g}$ must thus satisfy the boundary condition

$$\Phi_i(u, E, \xi_R) = 1 + (u - 1)\Theta(E - E_R) . \quad (4.4)$$

The generating function for e^+e^- annihilation at c.m. energy E_{cm} is that for two quarks of energy $E_{\text{cm}}/2$, each filling one hemisphere:

$$\Phi_{ee} = [\Phi_q(u, E_{\text{cm}}/2, 1)]^2 . \quad (4.5)$$

4.1 Next to double leading logarithms

For $\xi > \xi_R$ and $E > E_R$, we have to NDLA

$$\begin{aligned} \Phi_q(u, E, \xi) &= u\Delta_q(E, \xi) + \int_{\xi_R}^{\xi} \frac{d\xi'}{\xi'} \frac{\Delta_q(E, \xi)}{\Delta_q(E, \xi')} \int_{E_R/E}^1 dz \frac{\alpha_S(k_t^2)}{2\pi} P_{gq}(z) \Phi_q(u, E, \xi') \Phi_g(u, zE, \xi') , \\ \Phi_g(u, E, \xi) &= u\Delta_g(E, \xi) + \int_{\xi_R}^{\xi} \frac{d\xi'}{\xi'} \frac{\Delta_g(E, \xi)}{\Delta_g(E, \xi')} \int_{E_R/E}^1 dz \frac{\alpha_S(k_t^2)}{2\pi} \{ P_{gg}(z) \Phi_g(u, E, \xi') \Phi_g(u, zE, \xi') \\ &\quad + P_{qg}(z) [\Phi_q(u, E, \xi')]^2 \} . \end{aligned} \quad (4.6)$$

Using the expressions given earlier to eliminate the Sudakov factors, we may write these in the equivalent forms

$$\begin{aligned} \Phi_q(u, E, \xi) &= u + \int_{\xi_R}^{\xi} \frac{d\xi'}{\xi'} \int_{E_R/E}^1 dz \frac{\alpha_S(k_t^2)}{2\pi} P_{gq}(z) \Phi_q(u, E, \xi') [\Phi_g(u, zE, \xi') - 1] , \\ \Phi_g(u, E, \xi) &= u + \int_{\xi_R}^{\xi} \frac{d\xi'}{\xi'} \int_{E_R/E}^1 dz \frac{\alpha_S(k_t^2)}{2\pi} \{ P_{gg}(z) \Phi_g(u, E, \xi') [\Phi_g(u, zE, \xi') - 1] \\ &\quad + P_{qg}(z) [\{\Phi_q(u, E, \xi')\}^2 - \Phi_g(u, E, \xi')] \} . \end{aligned} \quad (4.7)$$

The solution for the quark generating function is then easily seen to be

$$\Phi_q(u, E, \xi) = u \exp \left\{ \int_{\xi_R}^{\xi} \frac{d\xi'}{\xi'} \int_{E_R/E}^1 dz \frac{\alpha_S(k_t^2)}{2\pi} P_{gq}(z) [\Phi_g(u, zE, \xi') - 1] \right\} . \quad (4.8)$$

We can solve for the gluon generating function by iteration, and then substitute in this equation to get the complete solution.

4.2 Leading double logarithms

In the DLA we keep only the singular parts of P_{gq} and P_{gg} and can drop P_{qq} . For brevity we define the logarithms as

$$\kappa = \ln(E/E_R) , \quad \lambda = \ln(\xi/\xi_R) . \quad (4.9)$$

Then

$$\Delta_{q,g}(\kappa, \lambda) = e^{-a_{q,g}\kappa\lambda} \quad (4.10)$$

where $a_{q,g} = C_{F,A}\bar{\alpha}_S$, and

$$\Phi_q(u, \kappa, \lambda) = u e^{-a_q\kappa\lambda} \exp \left\{ a_q \int_0^\kappa d\kappa' \int_0^\lambda d\lambda' \Phi_g(u, \kappa', \lambda') \right\} , \quad (4.11)$$

$$\Phi_g(u, \kappa, \lambda) = e^{-a_g\kappa\lambda} \left\{ u + a_g \int_0^\kappa d\kappa' \int_0^\lambda d\lambda' e^{a_g\kappa'\lambda'} \Phi_g(u, \kappa', \lambda') \Phi_g(u, \kappa', \lambda') \right\} .$$

We can simplify the equation for Φ_g by noting that

$$\frac{\partial}{\partial \lambda} (e^{a_g\kappa\lambda} \Phi_g) = a_g e^{a_g\kappa\lambda} \Phi_g \int_0^\kappa d\kappa' \Phi_g(u, \kappa', \lambda) \quad (4.12)$$

so that

$$\ln (e^{a_g\kappa\lambda} \Phi_g) = \int_0^\kappa d\kappa' \int_0^\lambda d\lambda' \Phi_g(u, \kappa', \lambda') + C(u, \kappa) \quad (4.13)$$

where the boundary condition $\Phi_g(u, \kappa, 0) = u$ for all $\kappa \geq 0$ implies $C(u, \kappa) = \ln u$. Thus

$$\Phi_g(u, \kappa, \lambda) = u e^{-a_g\kappa\lambda} \exp \left\{ a_g \int_0^\kappa d\kappa' \int_0^\lambda d\lambda' \Phi_g(u, \kappa', \lambda') \right\} . \quad (4.14)$$

Comparing with (4.11) we see that in the DLA

$$\Phi_q = u (\Phi_g/u)^{C_F/C_A} . \quad (4.15)$$

5. Jet rates

From the quark and gluon generating functions in the previous section it is relatively straightforward to construct the jet rates to next-to-double log accuracy (NDLA).

In terms of the logarithmic variables (4.9) the quark and gluon integrated splitting kernels (3.4) and (3.7) are simply

$$\mathcal{P}_q(E, \xi) = \bar{\alpha}_S \lambda C_F \left[\kappa - \frac{3}{4} \right] + \frac{1}{2} C_F b_0 \bar{\alpha}_S^2 \lambda \kappa [2\lambda + \kappa] , \quad (5.1)$$

$$\mathcal{P}_g(E, \xi) = \bar{\alpha}_S \lambda [C_A \kappa - b_0] + \frac{1}{2} C_A b_0 \bar{\alpha}_S^2 \lambda \kappa [2\lambda + \kappa] , \quad (5.2)$$

where

$$\bar{\alpha}_S \equiv \bar{\alpha}_S(\kappa, \lambda) = \frac{1}{b_0(2\kappa + \lambda + \mu)} \quad (5.3)$$

with $\mu = \ln(E_R^2 \xi_R / \Lambda^2)$.

The quark generating function is

$$\Phi_q(u, E, \lambda) = u \exp \left\{ C_F \int_0^\lambda d\lambda' \int_0^\kappa d\kappa' \left(\bar{\alpha}_S(\kappa', \lambda') - \frac{3}{4} \bar{\alpha}_S e^{\kappa' - \kappa} \right) [\Phi_g(u, \kappa', \lambda') - 1] \right\} . \quad (5.4)$$

In the fixed-order expansion it is not necessary to include the running coupling for the finite terms in the splitting functions, as doing so would affect results beyond the NDLA. In these finite term we can set terms proportional to $e^{-\kappa}$ to 0 after the integrations, which is equivalent to allowing the original z integration to range over $(0, 1)$, since it is not singular in energy. However, we have to bear in mind that the relevant integrals vanish when $\kappa = 0$.

Defining

$$\Gamma_g(\kappa', \lambda', \kappa) = C_A \left[\bar{\alpha}_S(\kappa', \lambda') - \frac{11}{12} \bar{\alpha}_S e^{\kappa' - \kappa} \right] , \quad (5.5)$$

$$\Gamma_q(\kappa', \lambda', \kappa) = C_F \left[\bar{\alpha}_S(\kappa', \lambda') - \frac{3}{4} \bar{\alpha}_S e^{\kappa' - \kappa} \right] , \quad (5.6)$$

we solve the gluon generating function by iteration to third order in u .

$$\Phi_g^{(1)}(u, \kappa, \lambda) = u \Delta_g(\kappa, \lambda) \quad (5.7)$$

$$\begin{aligned} \Phi_g^{(2)}(u, \kappa, \lambda) = u \Delta_g(\kappa, \lambda) & \left(1 + u \int_0^\lambda d\lambda' \int_0^\kappa d\kappa' \left\{ \Gamma_g(\kappa', \lambda', \kappa) \Delta_g(\kappa', \lambda') \right. \right. \\ & \left. \left. + \bar{\alpha}_S \frac{n_f}{6} \frac{\Delta_q^2(\kappa, \lambda')}{\Delta_g(\kappa, \lambda')} e^{\kappa' - \kappa} \right\} \right) \end{aligned} \quad (5.8)$$

$$\begin{aligned} \Phi_g^{(3)}(u, \kappa, \lambda) = u \Delta_g(\kappa, \lambda) & \left(1 + u \int_0^\lambda d\lambda' \int_0^\kappa d\kappa' \left\{ \Gamma_g(\kappa', \lambda', \kappa) \Delta_g(\kappa', \lambda') \right. \right. \\ & \times \left(1 + u \left[\int_0^\kappa + \int_0^{\kappa'} \right] d\kappa'' \int_0^{\lambda'} d\lambda'' \left\{ \Gamma_g(\kappa'', \lambda'', \bar{\kappa}) \Delta_g(\kappa'', \lambda'') \right. \right. \\ & \left. \left. + \bar{\alpha}_S \frac{n_f}{6} \frac{\Delta_q^2(\kappa', \lambda'')}{\Delta_g(\kappa', \lambda'')} e^{\kappa'' - \bar{\kappa}} \right\} \right) \\ & \left. + \bar{\alpha}_S \frac{n_f}{6} \frac{\Delta_q^2(\kappa, \lambda')}{\Delta_g(\kappa, \lambda')} e^{\kappa' - \kappa} \left(1 + 2u C_F \int_0^{\lambda'} d\lambda'' \int_0^\kappa d\kappa'' \right. \right. \\ & \left. \left. \times \Gamma_q(\kappa'', \lambda'', \kappa) \Delta_g(\kappa'', \lambda'') \right) \right\} \right). \end{aligned} \quad (5.9)$$

In the result for $\Phi_g^{(3)}$ we have defined $\bar{\kappa} = \kappa(\kappa')$ when the κ'' integral ranges from 0 to κ (κ'). Defining

$$\Psi_g^{(n)} = 2 \int_0^\kappa d\kappa' \int_0^\lambda d\lambda' \Gamma_q(\kappa', \lambda', \kappa) \Phi_g^{(n)}(u, \kappa', \lambda'), \quad (5.10)$$

the generating function for resummed rates up to 5 jets at NDLA is then

$$\Phi_{ee}^{(5)} = u^2 \Delta_q^2(\kappa, \lambda) \left\{ 1 + \Psi_g^{(3)} + \frac{1}{2} (\Psi_g^{(2)})^2 + \frac{1}{3!} (\Psi_g^{(1)})^3 \right\}. \quad (5.11)$$

Substituting in (5.4) and using (4.5), we find the rates for e^+e^- annihilation in the form

$$R_n^{ee} = \delta_{2,n} + \sum_{j \geq n-2} \bar{\alpha}_S^j (R_{n,2j} + R_{n,2j-1}) \quad (5.12)$$

for $n \leq 5$ and $j \leq 3$, where n is the number of resolved jets and $R_{n,i}$ has i powers of

the logarithms (either κ or λ):

$$\begin{aligned}
R_{22} &= -2C_F\kappa\lambda \\
R_{21} &= \frac{3}{2}C_F\lambda \\
R_{24} &= 2C_F^2\kappa^2\lambda^2 \\
R_{23} &= [-b_0(2\kappa + \lambda) - 3C_F\lambda] C_F\kappa\lambda \\
R_{26} &= -\frac{4}{3}C_F^3\kappa^3\lambda^3 \\
R_{25} &= [2b_0(2\kappa + \lambda) + 3C_F\lambda] C_F^2\kappa^2\lambda^2
\end{aligned} \tag{5.13}$$

$$\begin{aligned}
R_{32} &= 2C_F\kappa\lambda \\
R_{31} &= -\frac{3}{2}C_F\lambda \\
R_{34} &= \left[-4C_F - \frac{1}{2}C_A\right] C_F\kappa^2\lambda^2 \\
R_{33} &= \left[b_0(2\kappa + \lambda) + \left(\frac{5}{3}C_A + 6C_F - \frac{1}{6}n_f\right)\lambda\right] C_F\kappa\lambda \\
R_{36} &= \left[4C_F^2 + C_AC_F + \frac{1}{9}C_A^2\right] C_F\kappa^3\lambda^3 \\
R_{35} &= \left[\left(-\frac{1}{2}C_A - 4C_F\right)b_0(2\kappa + \lambda) + \right. \\
&\quad \left.\left(-\frac{5}{9}C_A^2 - \frac{49}{12}C_AC_F - 9C_F^2 + \frac{1}{18}C_An_f + \frac{1}{3}C_Fn_f\right)\lambda\right] C_F\kappa^2\lambda^2
\end{aligned} \tag{5.14}$$

$$\begin{aligned}
R_{44} &= \left[2C_F + \frac{1}{2}C_A\right] C_F\kappa^2\lambda^2 \\
R_{43} &= \left[-\frac{5}{3}C_A - 3C_F + \frac{1}{6}n_f\right] C_F\kappa\lambda^2 \\
R_{46} &= \left[-4C_F^2 - 2C_AC_F - \frac{5}{18}C_A^2\right] C_F\kappa^3\lambda^3 \\
R_{45} &= \left[\left(\frac{1}{2}C_A + 2C_F\right)b_0(2\kappa + \lambda) + \right. \\
&\quad \left.\left(\frac{37}{24}C_A^2 + \frac{49}{6}C_AC_F + 9C_F^2 - \frac{1}{9}C_An_f - \frac{13}{18}C_Fn_f\right)\lambda\right] C_F\kappa^2\lambda^2
\end{aligned} \tag{5.15}$$

$$\begin{aligned}
R_{56} &= \left[\frac{4}{3}C_F^2 + C_AC_F + \frac{1}{6}C_A^2\right] C_F\kappa^3\lambda^3 \\
R_{55} &= \left[-\frac{71}{72}C_A^2 - \frac{49}{12}C_AC_F - 3C_F^2 + \frac{1}{18}C_An_f + \frac{7}{18}C_Fn_f\right] C_F\kappa^2\lambda^3
\end{aligned} \tag{5.16}$$

One check on these results is that the DLA coefficients agree with the previous computation in [14] for R_{44} and R_{56} . A second is that the 2-jet inclusive fraction obtained by summing the rates is 1.

6. Average jet multiplicity

The average jet multiplicity, i.e. the mean number of resolved jets, as a function of the hard process scale Q , the angular resolution R and the minimum jet energy E_R , provides a useful overall measure of jet activity and substructure. As indicated by Eq. (4.3), this quantity is obtained simply from the first derivative of the relevant generating function.

6.1 Next to double leading logarithms

Writing the average jet multiplicity in e^+e^- annihilation as \mathcal{N}_{ee} , from the fixed-order jet rates (5.13)- (5.16) we obtain to $\mathcal{O}(\alpha_S^3)$

$$\begin{aligned} \mathcal{N}_{ee} = & 2 + \bar{\alpha}_S \left(2\kappa\lambda - \frac{3}{2}\lambda \right) C_F \\ & + \bar{\alpha}_S^2 \left(\frac{1}{2}C_A\kappa\lambda + 2b_0(2\kappa + \lambda) - \frac{5}{3}C_A\lambda + \frac{1}{6}n_f\lambda \right) C_F\kappa\lambda \\ & + \bar{\alpha}_S^3 \left(\frac{1}{18}C_A^2\kappa\lambda + \frac{1}{2}b_0C_A(2\kappa + \lambda) + \frac{1}{18}n_fC_F\lambda - \frac{31}{72}C_A^2\lambda \right) C_F\kappa^2\lambda^2 . \end{aligned} \quad (6.1)$$

Here the terms in b_0 originate from including the running coupling. We see that these terms enhance the average jet multiplicity with respect to a fixed-coupling calculation.

To perform an all-orders resummation to NDLA, we repeat the analysis of [15], for the inclusive algorithms instead of the exclusive k_t algorithm. In terms of the generating functions, we have

$$\mathcal{N}_{ee}(E, \xi) = \left. \frac{\partial \Phi_{ee}}{\partial u} \right|_{u=1} = 2 \left. \frac{\partial \Phi_q}{\partial u} \right|_{u=1} = 2\mathcal{N}_q \quad (6.2)$$

where $\mathcal{N}_{q,g}$, the average quark and gluon jet multiplicities, satisfy the equations

$$\mathcal{N}_q(E, \xi) = 1 + \int_{\xi_R}^{\xi} \frac{d\xi'}{\xi'} \int_{E_R/E}^1 dz \frac{\alpha_S(k_t^2)}{2\pi} P_{gq}(z) \mathcal{N}_g(zE, \xi') \quad (6.3)$$

$$\begin{aligned} \mathcal{N}_g(E, \xi) = & 1 + \int_{\xi_R}^{\xi} \frac{d\xi'}{\xi'} \int_{E_R/E}^1 dz \frac{\alpha_S(k_t^2)}{2\pi} \{ P_{gg}(z) \mathcal{N}_g(zE, \xi') \\ & + P_{qg}(z) [2\mathcal{N}_q(E, \xi') - \mathcal{N}_g(E, \xi')] \} . \end{aligned} \quad (6.4)$$

In Appendix A we show that (6.3) is equivalent to the following PDE in terms of the logarithmic variables (4.9):

$$\frac{\partial^2 \mathcal{N}_q}{\partial \kappa \partial \lambda} = C_F \bar{\alpha}_S \left(\mathcal{N}_g - \frac{3}{4} \frac{\partial \mathcal{N}_g}{\partial \kappa} \right) , \quad (6.5)$$

with boundary conditions $\mathcal{N}_q(\kappa, 0) = \mathcal{N}_q(0, \lambda) = 1$.

Similarly we find from (6.4)

$$\frac{\partial^2 \mathcal{N}_g}{\partial \kappa \partial \lambda} = \bar{\alpha}_S \left[C_A \mathcal{N}_g - \left(\frac{11}{12} C_A + \frac{n_f}{6} \right) \frac{\partial \mathcal{N}_g}{\partial \kappa} \right] + \frac{n_f}{3} \bar{\alpha}_S \frac{\partial \mathcal{N}_q}{\partial \kappa}, \quad (6.6)$$

where to the required accuracy we may set in the last term

$$\frac{\partial \mathcal{N}_q}{\partial \kappa} = \frac{C_F}{C_A} \frac{\partial \mathcal{N}_g}{\partial \kappa}, \quad (6.7)$$

so that finally

$$\frac{\partial^2 \mathcal{N}_g}{\partial \kappa \partial \lambda} = \bar{\alpha}_S \left[C_A \mathcal{N}_g - \left(\frac{11}{12} C_A + \frac{n_f}{6} - \frac{n_f C_F}{3 C_A} \right) \frac{\partial \mathcal{N}_g}{\partial \kappa} \right], \quad (6.8)$$

with boundary conditions $\mathcal{N}_g(\kappa, 0) = \mathcal{N}_g(0, \lambda) = 1$.

Note that the n_f dependence in (6.8) is very weak and vanishes in the large- N limit:

$$\frac{n_f}{6} - \frac{n_f C_F}{3 C_A} = \frac{n_f}{6 N^2} = \frac{n_f}{54}. \quad (6.9)$$

This is because at large N a $q\bar{q}$ pair from gluon splitting radiates like a gluon.

6.2 Leading double logarithms

Dropping the non-singular parts of the splitting functions, we have from (4.15)

$$\mathcal{N}_{ee} = 2\mathcal{N}_q = 2 + 2 \frac{C_F}{C_A} (\mathcal{N}_g - 1) \quad (6.10)$$

where

$$\frac{\partial^2 \mathcal{N}_g}{\partial \kappa \partial \lambda} = C_A \bar{\alpha}_S \mathcal{N}_g. \quad (6.11)$$

In the leading double log approximation, α_S is treated as a constant. Then the solution to (6.11) is a modified Bessel function:

$$\mathcal{N}_g(\kappa, \lambda) = \sum_{n=0}^{\infty} \frac{(C_A \bar{\alpha}_S \kappa \lambda)^n}{(n!)^2} = I_0 \left(2 \sqrt{C_A \bar{\alpha}_S \kappa \lambda} \right). \quad (6.12)$$

The asymptotic behaviour for large argument,

$$I_0(y) \sim \frac{e^y}{\sqrt{2\pi y}}, \quad (6.13)$$

implies that for high energy and small cone size

$$\begin{aligned} \mathcal{N}_{ee} \sim 2 \left(1 - \frac{C_F}{C_A} \right) + \frac{C_F}{\sqrt{\pi} C_A} \left[C_A \bar{\alpha}_S \ln \left(\frac{E_{\text{cm}}}{2E_R} \right) \ln \left(\frac{1}{\xi_R} \right) \right]^{-\frac{1}{4}} \times \\ \exp \left[2 \sqrt{C_A \bar{\alpha}_S \ln \left(\frac{E_{\text{cm}}}{2E_R} \right) \ln \left(\frac{1}{\xi_R} \right)} \right]. \end{aligned} \quad (6.14)$$

6.3 Running coupling

Taking into account the running of $\bar{\alpha}_S$ to next-to-leading order, we have

$$\frac{\partial^2}{\partial \kappa \partial \lambda} I_0 \left(2\sqrt{C_A \bar{\alpha}_S \kappa \lambda} \right) = [1 - b_0(2\kappa + \lambda)\bar{\alpha}_S + \mathcal{O}(\alpha_S^2)] I_0 \left(2\sqrt{C_A \bar{\alpha}_S \kappa \lambda} \right) . \quad (6.15)$$

Thus if we drop terms of relative order α_S^2 the solution to (6.11) is

$$N_g = [1 + b_0(2\kappa + \lambda)\bar{\alpha}_S] I_0 \left(2\sqrt{C_A \bar{\alpha}_S \kappa \lambda} \right) , \quad (6.16)$$

which agrees with the b_0 -dependent terms in (6.1). However, for large κ and/or λ , $b_0(2\kappa + \lambda)\bar{\alpha}_S \sim 1$ and therefore we need to take into account the running of α_S to all orders.

6.4 Numerical solution: DLA

Treating the running of α_S to all orders, as in Eq. (5.3), but still neglecting the finite parts of the splitting functions, we have in place of (6.11)

$$\frac{\partial^2 \mathcal{N}_g}{\partial \kappa \partial \lambda} = c_g \frac{\mathcal{N}_g}{(2\kappa + \lambda + \mu)} , \quad (6.17)$$

with $c_g = C_A/b_0$. This PDE is not straightforward to solve explicitly. Its properties are discussed in Appendix B. It may however be solved numerically by discretization. Writing

$$\begin{aligned} \mathcal{N}_q(\kappa = ma, \lambda = nb) &= f_{m,n} \\ \mathcal{N}_g(\kappa = ma, \lambda = nb) &= g_{m,n} , \end{aligned} \quad (6.18)$$

we have

$$\frac{\partial^2 \mathcal{N}_g}{\partial \kappa \partial \lambda} \approx \frac{1}{ab} [g_{m+1,n+1} - g_{m+1,n} - g_{m,n+1} + g_{m,n}] \quad (6.19)$$

and

$$\begin{aligned} \frac{c_g \mathcal{N}_g}{(2\kappa + \lambda + \mu)} &\approx \frac{c_g}{4} \left[\frac{g_{m+1,n+1}}{2(m+1)a + (n+1)b + \mu} + \frac{g_{m+1,n}}{2(m+1)a + nb + \mu} \right. \\ &\quad \left. + \frac{g_{m,n+1}}{2ma + (n+1)b + \mu} + \frac{g_{m,n}}{2ma + nb + \mu} \right] . \end{aligned} \quad (6.20)$$

Equating these expressions, one can solve iteratively for $g_{m+1,n+1}$ starting from the boundary values $g_{0,n} = g_{m,0} = 1$.

6.5 Numerical solution: NDLA

To include the finite parts of the splitting functions, we may write (6.5) and (6.8) with equivalent precision as

$$\frac{\partial^2 \mathcal{N}_{q,g}}{\partial \kappa \partial \lambda} = c_{q,g} \left(1 - d_{q,g} \frac{\partial}{\partial \kappa} \right) \frac{\mathcal{N}_g}{2\kappa + \lambda + \mu}, \quad (6.21)$$

where $c_{q,g} = C_{F,A}/b_0$ and

$$d_q = \frac{3}{4}, \quad d_g = \frac{11}{12} + \frac{n_f}{6N^3}. \quad (6.22)$$

The PDEs (6.21) can be solved numerically by a simple extension of the method outlined above. For the discretized κ -derivative, we use

$$\frac{\partial \mathcal{N}_g}{\partial \kappa} \approx \frac{1}{2a} [g_{m+1,n+1} + g_{m+1,n} - g_{m,n+1} - g_{m,n}]. \quad (6.23)$$

We can then write the right-hand side of (6.21) as

$$\begin{aligned} \frac{c_g}{4} & \left[\frac{(1 - \delta_g)g_{m+1,n+1}}{2(m+1)a + (n+1)b + \mu} + \frac{(1 - \delta_g)g_{m+1,n}}{2(m+1)a + nb + \mu} \right. \\ & \left. + \frac{(1 + \delta_g)g_{m,n+1}}{2ma + (n+1)b + \mu} + \frac{(1 + \delta_g)g_{m,n}}{2ma + nb + \mu} \right], \end{aligned} \quad (6.24)$$

where

$$\delta_g = \frac{2}{a} d_g = \frac{2}{a} \left(\frac{11}{12} + \frac{n_f}{6N^3} \right), \quad (6.25)$$

and equate this to (6.19).

Similarly, to obtain the quark jet multiplicity we write

$$\frac{\partial^2 \mathcal{N}_q}{\partial \kappa \partial \lambda} \approx \frac{1}{ab} [f_{m+1,n+1} - f_{m+1,n} - f_{m,n+1} + f_{m,n}], \quad (6.26)$$

equate this to (6.24) with c_g, δ_g replaced by

$$c_q = \frac{C_F}{b_0}, \quad \delta_q = \frac{3}{2a} \quad (6.27)$$

to obtain the discrete equivalent of (6.5), and solve iteratively for $f_{m+1,n+1}$ starting from the boundary values $f_{0,n} = f_{m,0} = 1$.

7. Monte Carlo results: e^+e^-

In Fig. 1 results for the average jet multiplicity from the SHERPA Monte Carlo are compared with the fixed-coupling DLA results (for fixed $\alpha_S = 0.118$). The average jet multiplicity is shown for $E_{\text{cm}} = 200$ TeV and $E_R = 10$ GeV as a function of the jet

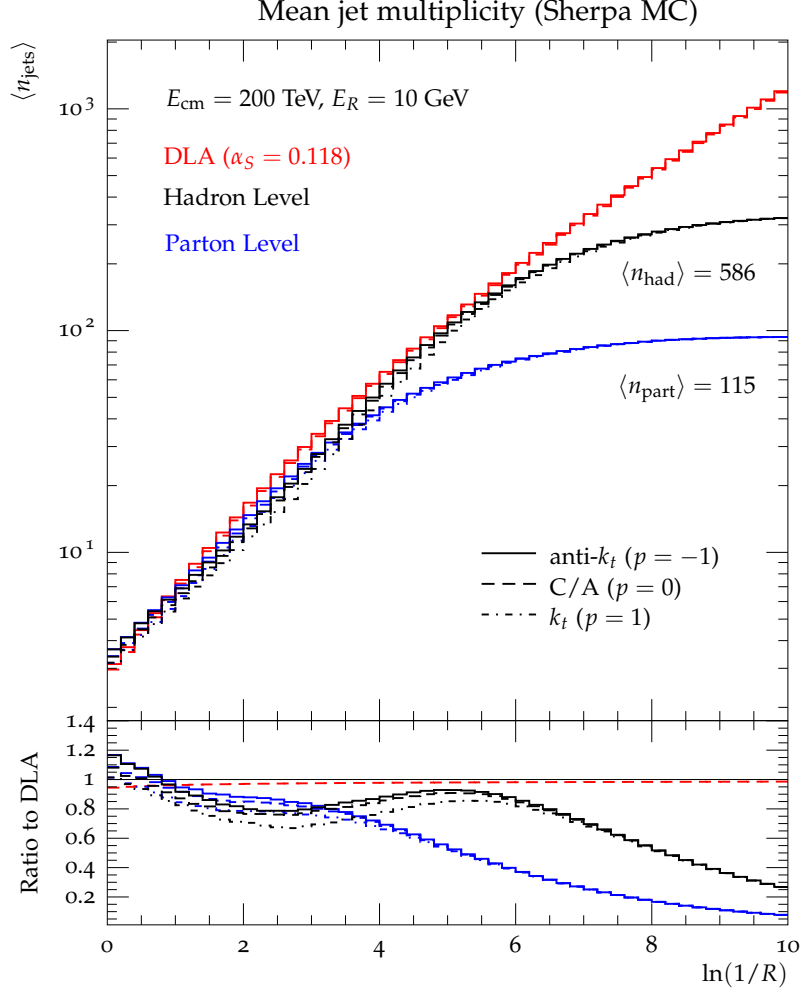


Figure 1: Average SHERPA jet multiplicity in e^+e^- annihilation at 200 TeV, compared to the DLA formula (6.12) (solid red), and the asymptotic limit (6.14) (red dashed). Black and blue show SHERPA hadron and parton levels, respectively. The solid, dashed and dot-dashed curves show anti- k_t , C/A and k_t algorithms, respectively.

radius parameter R . For these extreme values we expect the leading logarithms to dominate. At such a high energy, the asymptotic approximation (6.14) (red dashed) agrees well with the exact DLA formula (6.12) (solid red), even at $R \sim 1$. The Monte Carlo results follow the formula up to $\ln(1/R) \sim 4$, after which the parton-level Monte Carlo result saturates at the parton multiplicity (about 100 at the shower cutoff $Q_0 \sim 1$ GeV). Surprisingly, the hadron-level result follows the DLA perturbative prediction further, up to around 200 jets at $\ln(1/R) \sim 6$, before breaking away to saturate at the hadron multiplicity. We note that the quantitative agreement between Monte Carlo simulation, implementing the running coupling, and the analytic DLA result depends on the choice for the fixed coupling in the latter. For the extreme

c.m. energy we consider, a somewhat smaller value for α_s can improve the agreement with the simulation for larger values of $\ln(1/R)$. At the same time for lower values of $\ln(1/R)$, where we do not expect the DLA to fully apply, the agreement would get worse.

Figure 2 shows SHERPA results on the jet multiplicity as a function of c.m. energy, compared with the NDLA predictions for running α_s , derived numerically as explained in the previous Section. A small QCD scale $\Lambda \sim 50$ MeV in the NDLA formulae gives reasonable agreement with the Monte Carlo results, in which $\Lambda_{\overline{MS}} = 180$ MeV. Such a scale change is a next-to-NDLA effect and therefore within the uncertainties. At the smallest values of R , the Monte Carlo values are approaching the NDLA predictions from below, indicating further subleading effects. As in Fig. 1, the agreement at very small R is better at hadron level than at parton level.

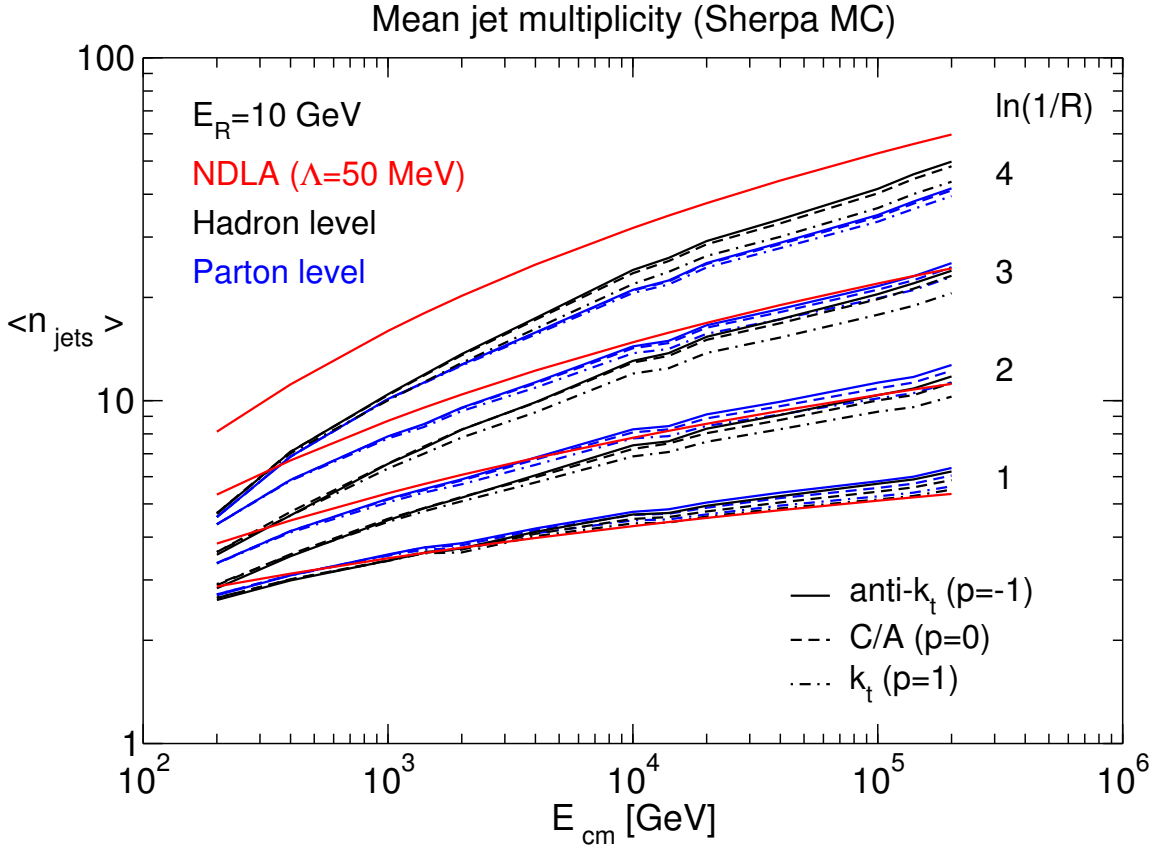


Figure 2: Average SHERPA jet multiplicity in e^+e^- annihilation as a function of centre-of-mass energy, compared to the NDLA result from Eqs. (6.5) and (6.8), shown in red. Black and blue show SHERPA hadron and parton levels, respectively. The solid, dashed and dot-dashed curves show anti- k_t , C/A and k_t algorithms, respectively.

It can be seen from both figures that the Monte Carlo results for all three inclusive algorithms, anti- k_t , C/A and k_t , are very similar, consistent with the lack

of dependence of our analytical predictions on the power p in Eqs. (2.1) and (2.2). The jet multiplicity is systematically slightly lower for the k_t algorithm, which we conjecture is due to its tendency to gather more low-momentum particles into jets from angles somewhat larger than the canonical jet radius R . This leads to a smaller number of jets for a given final-state multiplicity. This higher susceptibility of k_t jets to radiation was discussed in [16] where also a first-order estimate for the resulting effective area of k_t and C/A jets was derived. Qualitatively we observe the same behaviour as for the passive jet area discussed there: the effect increases for higher jet energies and larger nominal R .

8. Monte Carlo results: pp

The motivation for studying the inclusive family of jet algorithms is the similarity to phenomenologically relevant hadron collider algorithms. For jet rates, and therefore also the mean number of jets, the factorization of the PDF in the initial state holds at the double leading logarithmic order. In this section we compare the predictions of our generating functions with Monte Carlo calculations and determine the reliability of our resummation in the LHC context.

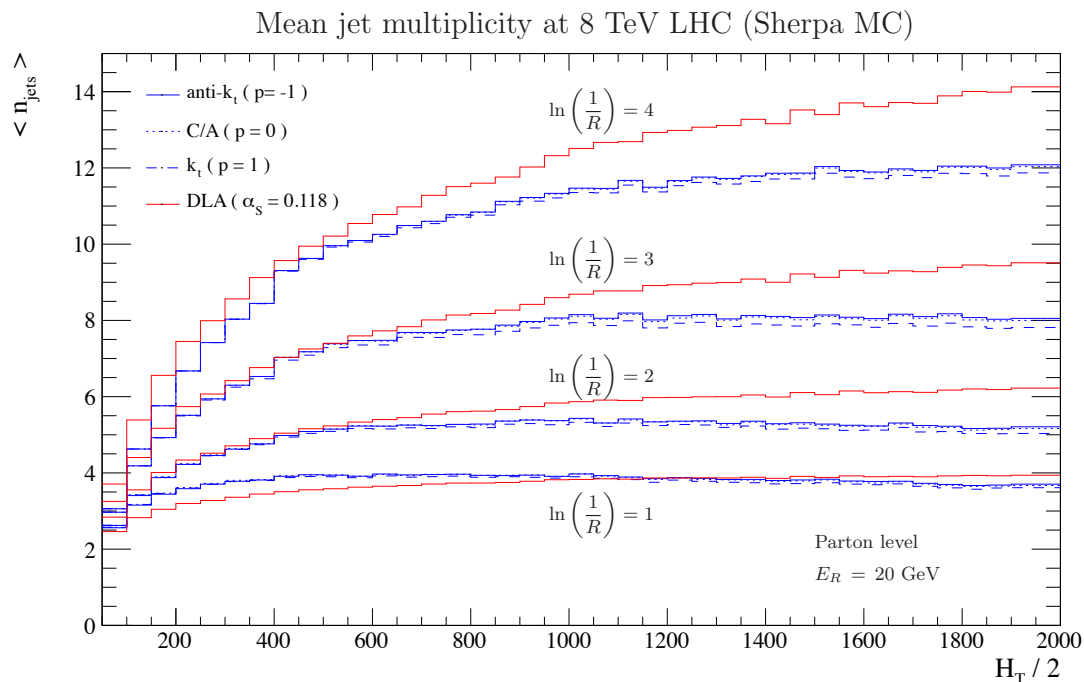


Figure 3: The proton-proton average jet multiplicity in inclusive di-jet production as a function of $H_T/2$. Compared is the DLA analytic prediction, shown in red, with parton shower results obtained with SHERPA. The solid, dashed and dot-dashed curves show anti- k_t , C/A and k_t algorithms, respectively.

Figure 3 shows results for pp collisions at the current LHC energy of 8 TeV. Here the average multiplicity of jets with $p_T > E_R = 20$ GeV is plotted as a function of $H_T/2$, where

$$H_T = \sum_i p_{T,i} , \quad (8.1)$$

the scalar sum of all jet transverse momenta. This choice corresponds closely with the initial parton transverse momenta in the QCD $2 \rightarrow 2$ hard scattering. Each point in the analytic result is weighted by the partonic fraction of final-state quarks versus gluons. In other words, the analytic result is

$$\langle n_{jj} \rangle = 2(c_q \langle n_q \rangle + c_g \langle n_g \rangle) , \quad (8.2)$$

where $c_q + c_g = 1$ and are determined from the proportions of final-state jets in the contributing $2 \rightarrow 2$ hard subprocesses. The considered values for R are the same as in Fig. 2. The multiplicity levels off and even decreases at high values of the primary parton p_T , owing to the transition from gluon to quark jets. Overall we observe a good agreement of the analytic estimate with the SHERPA simulation. As before we use a fixed coupling of $\alpha_s = 0.118$ in the DLA calculation. In particular for large values of $H_T/2$ a reduced value could be more appropriate, improving the agreement with simulation. Notably, the dependence on the jet algorithm for the simulated results is rather mild. As observed for e^+e^- collisions in the above, the k_t algorithm tends to produce slightly less jets for a given final-state multiplicity.

9. Jet scaling patterns

The work in [17–19] discussed the potential for extrapolating jet rates based on universal scaling patterns. These patterns are most easily classified in terms of the ratio of exclusive jet rates $R_{(n+1)/n} = \sigma_{n+1}/\sigma_n$. In the Durham (exclusive k_t) algorithm, it was found that for low multiplicity $n \leq \langle n_{jets} \rangle$, emissions are essentially Poisson-like so that $R_{(n+1)/n} \sim (n+1)^{-1}$. The tail of the multiplicity distribution then produces dominantly staircase or geometric scaling where $R_{(n+1)/n} \sim \text{constant}$. This regime is driven by the fractal nature of QCD radiation in the gluon dominated limit.

We know from previous work that the expected scaling patterns of jets can depend dramatically on the jet algorithm. One example of this is the JADE algorithm, where the non-exponentiation of the primary emissions precludes the Poisson extrapolation even in the pseudo-abelian limit [20]. In this section we would like to address scaling in the inclusive generalized k_t class of algorithms. This extends the results in [17] and strengthens the case for investigation at hadron colliders.

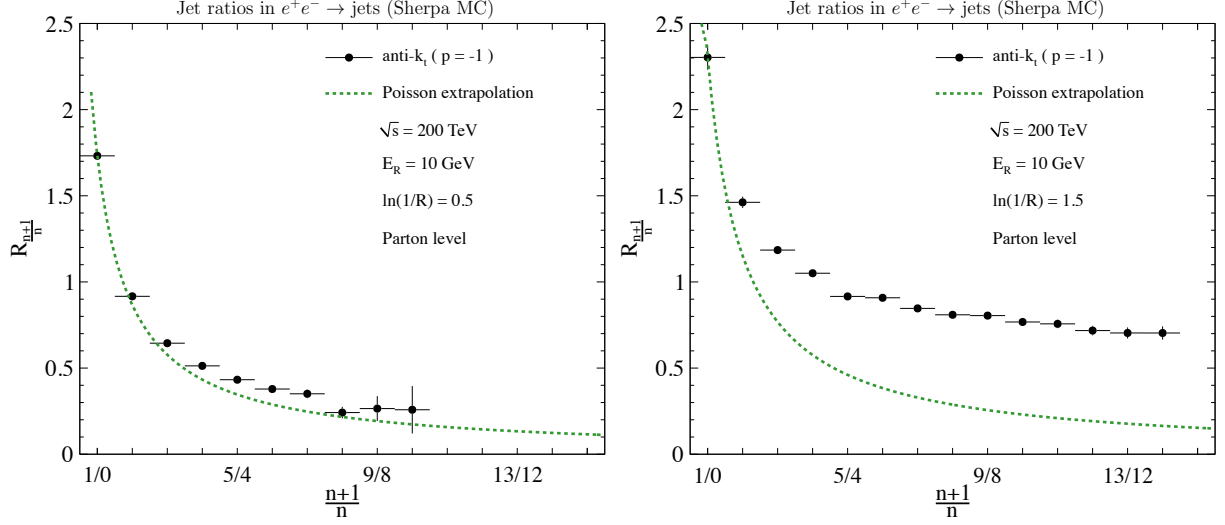


Figure 4: Jet ratios from SHERPA compared with the Poisson extrapolation from the first bin. For smaller jet sizes the Poisson behaviour breaks down as discussed in the text. Jet sizes here correspond to approximate R values of 0.61 (left) and 0.22 (right).

9.1 Poisson breaking components

With the leading logarithmic coefficients from Eqs. (5.13)-(5.16) it is easy to make some first statements about scaling in the generalized algorithm. It is clear from the structure of the coefficients that in the limit $C_A \rightarrow 0$ a perfect Poisson distribution emerges. Now a simple comparison between the generalized and Durham algorithms is the relative size of the Poisson breaking components in the lower multiplicity rates, for example the 2-gluon correlated emission contribution to the 4-jet rate (5.15). For the double-leading logarithmic coefficients to the 4-jet rate, R_{44} , we find $C_{44}^{\text{Durham}} \sim 2C_F^2 + (1/3)C_A C_F$ and $C_{44}^{\text{generalized}} \sim 2C_F^2 + (1/2)C_A C_F$ using the normalization of this work. The lowest-order Poisson breaking term is relatively larger in the generalized class of algorithms. We would thus expect the onset of staircase (geometric) scaling to come about even for smaller values of the logarithm, and to better match the staircase behaviour at lower multiplicity. Evidence of this may be found in the $Z + \text{jets}$ analysis presented in [17]. We present in table 1 the relative sizes of the Poisson breaking terms in the DLA expanding coefficients for the two algorithms compared with the idealized Poisson and staircase predictions. In this table we compute the 6-jet rate in the generalized algorithm by expanding the resummed jet rates from Eq. (5.11) to $\mathcal{O}(\alpha_s^4)$ which entirely determines the DLA resolved component.

	Poisson	Generalized k_t	Durham	Staircase
$R_{4/3}/R_{3/2}$	0.50	0.781	0.688	1
$R_{5/4}/R_{4/3}$	0.67	0.906	0.868	1
$R_{6/5}/R_{5/4}$	0.75	0.923	0.932	1

Table 1: Ratio of successive ratios for the generalized k_t and Durham jet rates compared with the idealized Poisson and staircase expectations.

9.2 R -dependence of scaling

A second question we wish to answer in this section is how the idealized scaling patterns depend on the jet radius parameter R in the generalized k_t algorithm. The additional handle provided by the separation of the angular and energy regulation allows us to probe an effect unbeknownst in the Durham algorithm. At the double leading logarithmic level, the resummed rates in the general algorithm are invariant under exchange of $\kappa \leftrightarrow \lambda$. Decreasing the size of the jet merely increases the overall logarithm.

In the simulation however, although we increase the overall logarithm when we require smaller resolved jets, we also change the relative contributions of the primary and secondary contributions due to kinematics. This effect is present even at the level of the ordered 2-gluon emission matrix element [21].

Using the parton shower we find that larger jet sizes dramatically increases the goodness of the fit with respect to the Poisson hypothesis. This has a clear interpretation. Correlated emissions in these events are predominantly intra-jet evolution for large jet radius. In addition, this effect is larger than that due to the breaking of $\kappa \leftrightarrow \lambda$ symmetry at NDLA.

To analyze this further we investigate the observable θ_{34} , defined as

$$\cos \theta_{34} \equiv \frac{\mathbf{p}_3 \cdot \mathbf{p}_4}{|\mathbf{p}_3||\mathbf{p}_4|}, \quad (9.1)$$

in reconstructed 4-jet events. The third and fourth hardest jets overwhelmingly originate from emitted gluons. As we suspect that in the simulation secondary Poisson breaking emissions are on average closer in θ_{34} , this observable gives another estimate on the size of these breaking effects. Comparing the prediction of pseudo-abelian QCD (in practice the parton shower with the $g \rightarrow gg$ and $g \rightarrow q\bar{q}$ branching off) we see in Fig. 5 that the jet size has a large effect on the relative size of the correlated emission component. In fact it appears that roughly $R \sim 0.4 - 0.5$ is the smallest value of the radius where the primary emissions are still the dominant contribution to the 4-jet rate.

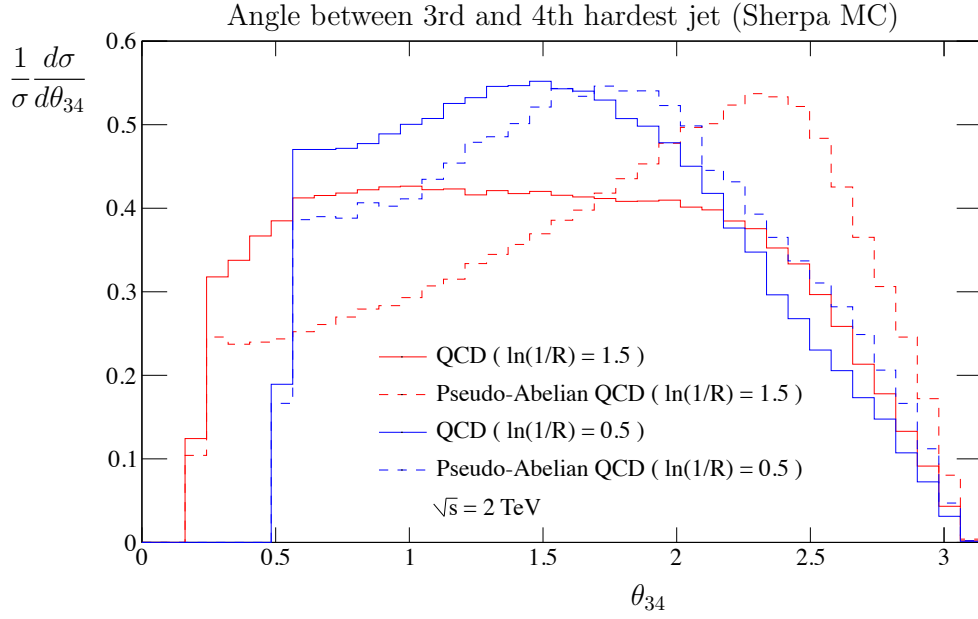


Figure 5: Differential cross-section with respect to the angle θ_{34} in exclusive 4-jet events. The solid curves in both cases represents the parton shower while the dashed curves have the gluon splitting function turned off (both to quarks and gluons). The large contribution of subsequent splittings between the red curves (small jet sizes) prevents a valid Poisson extrapolation as seen in Fig. 4.

10. Conclusions

In the present paper we have derived the generating functions of jet rates for the inclusive generalized k_t jet algorithms, valid in the next-to-double log approximation (NDLA), and used them to compute jet rates and average jet multiplicities as functions of the jet radius parameter R and the minimum jet energy E_R . At this level of precision, the results are independent of the power p that distinguishes the inclusive k_t , Cambridge/Aachen and anti- k_t algorithms.

The analytical results on e^+e^- annihilation are in broad agreement with those from the SHERPA Monte Carlo event generator, including the weak p -dependence. Surprisingly, the hadron-level Monte Carlo results follow the analytical predictions down to very small jet radii, well beyond the range of the perturbative parton shower, indicating that the cluster hadronization model in SHERPA is smoothly matched to perturbation theory. Analytical predictions for pp collisions at the current LHC energy also agree fairly well with SHERPA Monte Carlo results.

The SHERPA Monte Carlo results and generating function in the generalized k_t algorithm were also used to study the transition from Poisson-like to staircase-like behaviour in the $(n+1)/n$ jet ratios. The relatively larger non-abelian terms in the DLA e^+e^- jet rates, compared to the Durham algorithm, indicate a transition at

lower values of n . Furthermore, the jet radius was found to have a significant impact on scaling, where smaller jet sizes receive larger secondary contributions. This feature may be relevant for jet activity studies in jet substructure analyses.

These results represent the most detailed comparison to date between analytical and Monte Carlo predictions for the inclusive jet algorithms in current use at the LHC. Information about the dependence of jet rates and multiplicities at high jet energies and relatively small radii are particularly important for studies of boosted jets and jet substructure, which play an increasingly important role in searches for new physics at the LHC. Our leading-logarithmic predictions are relevant to such studies, and their fairly good agreement with Monte Carlo results is encouraging. Possible improvements such as optimized scale choices and empirical subleading contributions would be worth exploring. Studies of other features of the jet multiplicity distribution along these lines could also be useful.

Acknowledgements

BW acknowledges the support of a Leverhulme Trust Emeritus Fellowship, and thanks the Pauli Institute at ETH/University of Zurich for hospitality and support during part of this work. EG and SS acknowledge support by the Bundesministerium für Bildung und Forschung under contract 05H2012.

A. Derivation of the PDE for the average jet multiplicity

Differentiating Eq. (6.3) with respect to ξ and using (3.3), we have to NDLA

$$\begin{aligned} \xi \frac{\partial}{\partial \xi} \frac{\mathcal{N}_q - 1}{\bar{\alpha}_S} &= C_F \int_{E_R/E}^1 dz \left(\frac{1}{z} - \frac{3}{4} - b_0 \bar{\alpha}_S \frac{2}{z} \ln z \right) \mathcal{N}_g(zE, \xi) \\ &\quad + C_F b_0 \bar{\alpha}_S \int_{\xi_R}^{\xi} \frac{d\xi'}{\xi'} \int_{E_R/E}^1 \frac{dz}{z} \mathcal{N}_g(zE, \xi') . \end{aligned} \quad (\text{A.1})$$

But at the same level of precision we have (remembering the ξ dependence of $\bar{\alpha}_S$)

$$\xi \frac{\partial}{\partial \xi} \frac{\mathcal{N}_q - 1}{\bar{\alpha}_S} = \frac{1}{\bar{\alpha}_S} \xi \frac{\partial \mathcal{N}_q}{\partial \xi} + b_0 (\mathcal{N}_q - 1) \quad (\text{A.2})$$

and

$$C_F \bar{\alpha}_S \int_{\xi_R}^{\xi} \frac{d\xi'}{\xi'} \int_{E_R/E}^1 \frac{dz}{z} \mathcal{N}_g(zE, \xi') = \mathcal{N}_q - 1 , \quad (\text{A.3})$$

so that

$$\xi \frac{\partial \mathcal{N}_q}{\partial \xi} = C_F \bar{\alpha}_S \int_{E_R/E}^1 dz \left(\frac{1}{z} - \frac{3}{4} - b_0 \bar{\alpha}_S \frac{2}{z} \ln z \right) \mathcal{N}_g(zE, \xi) . \quad (\text{A.4})$$

Differentiating with respect to E and using (3.3) again,

$$E \frac{\partial}{\partial E} \frac{\xi}{\bar{\alpha}_S} \frac{\partial \mathcal{N}_g}{\partial \xi} = C_F \mathcal{N}_g + \frac{3}{4} C_F \int_{E_R/E}^1 dz [\mathcal{N}_g(zE, \xi) - \mathcal{N}_g(E, \xi)] \\ + 2C_F b_0 \bar{\alpha}_S \int_{E_R/E}^1 \frac{dz}{z} \mathcal{N}_g(zE, \xi). \quad (\text{A.5})$$

Now to the required precision

$$\int_{E_R/E}^1 dz [\mathcal{N}_g(zE, \xi) - \mathcal{N}_g(E, \xi)] = E \frac{\partial \mathcal{N}_g}{\partial E} \int_0^1 dz \ln z = -E \frac{\partial \mathcal{N}_g}{\partial E} \quad (\text{A.6})$$

and

$$2C_F b_0 \bar{\alpha}_S \int_{E_R/E}^1 \frac{dz}{z} \mathcal{N}_g(zE, \xi) = 2b_0 \xi \frac{\partial \mathcal{N}_g}{\partial \xi} \quad (\text{A.7})$$

while (remembering the E dependence of $\bar{\alpha}_S$)

$$E \frac{\partial}{\partial E} \frac{\xi}{\bar{\alpha}_S} \frac{\partial \mathcal{N}_g}{\partial \xi} = \frac{1}{\bar{\alpha}_S} E \frac{\partial}{\partial E} \xi \frac{\partial \mathcal{N}_g}{\partial \xi} + 2b_0 \xi \frac{\partial \mathcal{N}_g}{\partial \xi}. \quad (\text{A.8})$$

Thus we get the PDE

$$E \frac{\partial}{\partial E} \xi \frac{\partial \mathcal{N}_g}{\partial \xi} = C_F \bar{\alpha}_S \left(\mathcal{N}_g - \frac{3}{4} E \frac{\partial \mathcal{N}_g}{\partial E} \right), \quad (\text{A.9})$$

which gives (6.5) in terms of the logarithmic variables (4.9).

B. Properties of the PDE for the average jet multiplicity

The PDE (6.17) is not so straightforward to solve explicitly. If we consider changing to variables corresponding to the sum and difference of κ and λ , we find that the PDE is separable (one obtains ODEs that are equivalent to those describing the classical mechanics of the harmonic oscillator and the quantum mechanics of the Coulomb potential), but that the boundary conditions are not. As a result, one may prove that one cannot express the boundary conditions as a Fourier-Bessel series of orthogonal functions of Sturm-Liouville type.

On reflection, seeking an explicit solution in this way is perhaps ambitious, given that there is no guarantee that solutions to such a PDE can be written in terms of special functions. We thus proceed to an analysis of a rather different nature. This analysis will enable us to (i) establish that a solution exists; (ii) give a variety of infinite monotonic series of upper and lower bounds on that solution; and (iii) provide explicit series solutions that enable, *e.g.*, the asymptotic behaviour of the solution to be found in simple, closed form.

To proceed, it is useful to define $x = \frac{c_g}{2}(2\kappa + \frac{\mu}{2})$, $y = \frac{c_g}{2}(\lambda + \frac{\mu}{2})$, in terms of which

$$\mathcal{N}_{xy} = \frac{\mathcal{N}}{x+y}. \quad (\text{B.1})$$

We may recast the PDE in the integral forms

$$\frac{\partial \mathcal{N}}{\partial y} = \int_a^x dx' \frac{\mathcal{N}}{x' + y'}, \quad (\text{B.2})$$

where $a = \frac{c_g \mu}{4}$, or

$$\mathcal{N} = 1 + \int_a^y dy' \int_a^x dx' \frac{\mathcal{N}}{x' + y'}. \quad (\text{B.3})$$

The first of these forms, by the way, makes clear that a physically-acceptable solution, \mathcal{N} , if it exists, is a monotonically increasing function of $x, \forall y$, and of $y, \forall x$. (Proof: The mean number of jets must be ≥ 0 . Thus $\mathcal{N}_y > 0$; $\mathcal{N}_x > 0$ follows by symmetry in $x \leftrightarrow y$.)

The basic idea behind our analysis will be to first identify related PDEs for which we can find explicit solutions whose properties (such as their existence, continuity, $\mathcal{E}c$.) can be checked ‘by hand’. We then use these solutions as a crutch to derive properties (for example, the existence) of the solution of the original PDE.

Our first lemma is as follows. Suppose $\mathcal{M}^0(x, y)$ is a solution of the PDE

$$\mathcal{M}_{xy}^0 = \alpha(x, y) \mathcal{M}^0 \quad (\text{B.4})$$

subject to the same boundary conditions as \mathcal{N} . Suppose furthermore that $\alpha(x, y) \geq \frac{1}{x+y}$ almost everywhere¹ in the domain $D \equiv \{(x, y) | x, y \geq a\}$. Then

$$\mathcal{M}^1(x, y) \equiv 1 + \int_a^y dy' \int_a^x dx' \frac{\mathcal{M}^0}{x' + y'} \leq \mathcal{M}^0, \quad \forall (x, y) \in D. \quad (\text{B.5})$$

(For brevity, we denote a double integral of the type that appears on the RHS by $\int \frac{\mathcal{M}^0}{x+y}$ henceforth.) Proof of this follows immediately from the fact that $1 + \int \frac{\mathcal{M}^0}{x+y} \leq 1 + \int \alpha \mathcal{M}^0 = \mathcal{M}^0$ everywhere in D . Continuing in this vein, we define an infinite sequence whose elements are

$$\mathcal{M}^{i+1}(x, y) \equiv 1 + \int \frac{\mathcal{M}^i}{x+y}. \quad (\text{B.6})$$

Now

$$\mathcal{M}^{i+1} - \mathcal{M}^i = \int \frac{\mathcal{M}^i - \mathcal{M}^{i-1}}{x+y} \quad (\text{B.7})$$

so $\mathcal{M}^i - \mathcal{M}^{i-1} \leq 0 \quad \forall (x, y) \in D \implies \mathcal{M}^{i+1} - \mathcal{M}^i \leq 0 \quad \forall (x, y) \in D$ and proof that this sequence is monotonically decreasing follows by induction, since we have already proven that $\mathcal{M}^1 \leq \mathcal{M}^0$ everywhere in D .

¹We use the usual language of measure theory.

We next prove that the sequence $\{\mathcal{M}^i\}$ is bounded below by zero, if $\mathcal{M}^0 > 0$ (which we can verify by hand given an explicit \mathcal{M}^0), and hence converges. To do so, we simply note that positivity of \mathcal{M}^0 implies positivity of $\mathcal{M}^i \forall i$ by induction, given (B.6) and $\mathcal{M}^0 > 0$.

We now prove that the sequence $\{\mathcal{M}^i\}$ converges to a solution, $\overline{\mathcal{N}}$ of the original PDE. Indeed

$$\overline{\mathcal{N}} \equiv \lim_{i \rightarrow \infty} \mathcal{M}^i = \lim_{i \rightarrow \infty} \mathcal{M}^{i+1} = 1 + \lim_{i \rightarrow \infty} \int \frac{\mathcal{M}^i}{x+y} = 1 + \int \lim_{i \rightarrow \infty} \frac{\mathcal{M}^i}{x+y} = 1 + \int \frac{\overline{\mathcal{N}}}{x+y}, \quad (\text{B.8})$$

where, in the penultimate step we used the Theorem of Monotone Convergence of Lebesgue integration.²

We shall call $\overline{\mathcal{N}}$ a *supersolution* of (B.1). Evidently, given any suitable α and \mathcal{M}^0 , we can use the above results to prove that a solution exists and to find an infinite series of monotonically decreasing upper bounds on it. This raises the question of uniqueness, however: different starting points, (α, \mathcal{M}^0) , may lead to different supersolutions.

It is straightforward to prove the following weak version of uniqueness, for sequences that are ‘nested’ in the following sense. Suppose two sequences $\{\mathcal{M}^i\}$ and $\{\mathcal{M}'^i\}$ converge on supersolutions $\overline{\mathcal{N}}$ and $\overline{\mathcal{N}}'$, respectively. If $\exists i, j, k, l$ such that elements $\mathcal{M}^i \geq \mathcal{M}'^j$ and $\mathcal{M}^k \leq \mathcal{M}'^l$ almost everywhere, then $\overline{\mathcal{N}} = \overline{\mathcal{N}}'$. Proof: $\mathcal{M}^i \geq \mathcal{M}'^j \implies \mathcal{M}^i \geq \overline{\mathcal{N}}'$. But $\mathcal{M}^i - \overline{\mathcal{N}}' \geq 0 \implies \int \frac{\mathcal{M}^i - \overline{\mathcal{N}}'}{x+y} = \mathcal{M}^{i+1} - \overline{\mathcal{N}}' \geq 0 \implies \overline{\mathcal{N}} \geq \overline{\mathcal{N}}'$. Similarly, $\mathcal{M}^k \leq \mathcal{M}'^l \implies \overline{\mathcal{N}} \leq \overline{\mathcal{N}}' \implies \overline{\mathcal{N}} = \overline{\mathcal{N}}'$, QED. Unfortunately, this weak version of uniqueness does not even imply that any two supersolutions coincide, since the sequences that lead to them may be wholly contained in distinct intervals, or, even if they are nested, an element of one sequence need not exceed an element of the other almost everywhere. But this weak version is still useful where it applies, in that it may give us a more rapidly converging sequence of upper bounds on a solution.

Let us now discuss lower bounds. Similar arguments to those just given show that if we instead started with a solution, $\mathcal{L}^0(x, y)$ to a PDE of the form

$$\mathcal{L}_{xy}^0 = \beta(x, y) \mathcal{L}^0, \quad (\text{B.9})$$

with $\beta < \frac{1}{x+y}$ almost everywhere on D , then we shall obtain an infinite, monotonically *increasing* sequence. We can, furthermore, prove that this sequence is bounded above and hence converges, provided we can show by hand that $\mathcal{L}^0 \leq \mathcal{M}^0$, for some \mathcal{M}^0

²This Theorem requires that the \mathcal{M}^i be measurable, but this follows from the fact \mathcal{M}^0 , and hence \mathcal{M}^i are C^0 . Note, however, that we have not proven that $\overline{\mathcal{N}}$ itself is C^0 , let alone C^1 or C^2 . Strictly speaking, therefore, we have proven that $\overline{\mathcal{N}}$ is a solution of the integral equation, which is what we started with, rather than the PDE. We shall in any case ignore this subtlety in what follows.

that converges to a supersolution $\overline{\mathcal{N}}$. To wit, suppose we have shown that $\mathcal{L}^0 \leq \mathcal{M}^0$. It then follows by induction that $\mathcal{L}^i \leq \mathcal{M}^i$, since $\mathcal{L}^{i+1} - \mathcal{M}^{i+1} = \int \frac{\mathcal{L}^i - \mathcal{M}^i}{x+y}$, and thus that $\lim_{i \rightarrow \infty} \mathcal{L}^i \leq \overline{\mathcal{N}}$. Again, by the Theorem of Monotone Convergence, $\{\mathcal{L}^i\}$ converges to a solution of (B.1) that we call a *subsolution* and denote by $\underline{\mathcal{N}}$. By a straightforward generalization of the result just proven, a subsolution is less than or equal to any supersolution for which it can be shown that any one element in the sequence defining the former is less than or equal to (almost everywhere) any element in the sequence defining the latter.

Yet again, this result is insufficient to establish uniqueness, but it does provide a way to obtain both upper and lower bounds on a given solution. Obviously, if we can prove (or assume) uniqueness, then all of the aforementioned super- and sub-solutions coincide, meaning that any of the aforementioned sequences can be used to bound the solution to arbitrary, known precision.

We can prove uniqueness of the solution heuristically out to any finite (x, y) by explicit construction. Given $\mathcal{N} = 1$ on the boundary, we divide the intervals $[a, x]$ and $[a, y]$ into regions of size δx and δy and find the unique solution $\mathcal{N}(a + \delta x, y) \simeq 1 + \delta x \int_a^y \frac{dy'}{a+y'}$, with a similar result for $\mathcal{N}(x, a + \delta y)$. We then extrapolate to the region $[a + \delta x, a + 2\delta x]$ and so on. Given that we have proven the existence of the solution, we do not need to worry that the extrapolations in the x and y directions might not coincide. To make this proof rigorous outside a neighbourhood of the boundary, we would then have to take the limit $\delta x, \delta y \rightarrow 0$. We will content ourselves with assuming that the solution is unique.

As we stressed above, all of our results on super- and sub-solutions are contingent on showing the existence of solutions to the equations (B.4) and (B.9), for suitable α and β , satisfying the BCs, and showing that they have the desired properties. We do this by supplying explicit solutions for functions α, β of the combination $(x-a)(y-a)$, for which the PDE reduces to an ODE, and for which the boundary conditions reduce to a single boundary condition at $(x-a)(y-a) = 0$.

To obtain lower bounds, one may start, for example, with the solution $\mathcal{L}^0 = 1$ of the PDE $\mathcal{L}_{xy}^0 = 0$, since $\frac{1}{x+y} > 0$ everywhere in D . We thus obtain

$$\mathcal{L}^1 \geq 1 + (x+y) \ln(x+y) - (x+a) \ln(x+a) - (a+y) \ln(a+y) + 2a \ln 2a, \quad (\text{B.10})$$

with subsequent \mathcal{L}^i given in terms of polylogarithms. At fixed y and large x , for example, we deduce that $\mathcal{L}^1 \sim (y-a) \ln x$. More generally, $\mathcal{L}^n \sim \frac{((y-a) \ln x)^n}{(n!)^2} \implies \mathcal{N} \sim I_0(y-a) \ln x \sim \frac{x^{(y-a)}}{\sqrt{2\pi(y-a) \log x}}$.

To find an upper bound, we can use $\alpha = \frac{1}{2a}$ which evidently exceeds $\frac{1}{x+y}$ throughout D . This yields an initial upper bound of the form

$$\mathcal{M}^0 = I_0(\sqrt{2a(x-a)(y-a)}). \quad (\text{B.11})$$

A more stringent upper bound may be obtained from $\alpha = \frac{1}{2\sqrt{(x-a)(y-a)}} > \frac{1}{x+y}$, with solution

$$\mathcal{M}^0 = I_0(2\sqrt{2}(x-a)^{\frac{1}{4}}(y-a)^{\frac{1}{4}}). \quad (\text{B.12})$$

For a better lower bound, we can, provided $a > 1$, solve

$$\mathcal{L}_{xy}^0 = \frac{2}{(x+y)^2} \mathcal{L}^0, \quad (\text{B.13})$$

whose solution is

$$\mathcal{L}^0 = \frac{xy + a^2}{a(x+y)}. \quad (\text{B.14})$$

It is soothing to verify explicitly that the resulting subsolution has the same asymptotic behaviour, at large x and fixed y , as the subsolution starting from $L^0 = 1$ derived above. Indeed, in this limit, $\mathcal{L}^0 \sim \frac{y}{a}$, $\mathcal{L}^1 \sim \frac{y^2 - a^2}{2a} \log x$, and so on, with $\mathcal{L}^n \sim \frac{(y-a)^n (y+na)}{a(n+1)!} \frac{\log x}{n!}$. At large n , we thus find $\mathcal{L}^n \sim \frac{(y-a)^n}{n!} \frac{\log x}{n!}$, exactly as before.

References

- [1] G. P. Salam, “Towards Jetography,” *Eur. Phys. J. C* **67** (2010) 637 [arXiv:0906.1833 [hep-ph]].
- [2] S. Catani, Y. L. Dokshitzer, M. Olsson, G. Turnock and B. R. Webber, “New clustering algorithm for multi-jet cross sections in e^+e^- annihilation,” *Phys. Lett. B* **269** (1991) 432.
- [3] S. Catani, Y. L. Dokshitzer, M. H. Seymour and B. R. Webber, “Longitudinally invariant k_t clustering algorithms for hadron hadron collisions,” *Nucl. Phys. B* **406** (1993) 187.
- [4] S. D. Ellis and D. E. Soper, “Successive combination jet algorithm for hadron collisions,” *Phys. Rev. D* **48** (1993) 3160 [hep-ph/9305266].
- [5] M. Cacciari, G. P. Salam and G. Soyez, “The anti- k_t jet clustering algorithm,” *JHEP* **0804** (2008) 063 [arXiv:0802.1189 [hep-ph]].
- [6] M. Cacciari, G. P. Salam and G. Soyez, “FastJet user manual,” *Eur. Phys. J. C* **72** (2012) 1896 [arXiv:1111.6097 [hep-ph]].
- [7] K. Konishi, A. Ukawa and G. Veneziano, “Jet Calculus: A Simple Algorithm for Resolving QCD Jets,” *Nucl. Phys. B* **157** (1979) 45.
- [8] Y. L. Dokshitzer, V. A. Khoze, A. H. Mueller and S. I. Troian, “Basics of perturbative QCD,” Gif-sur-Yvette, France: Ed. Frontieres (1991) 274 p. (Basics of)
- [9] R. K. Ellis, W. J. Stirling and B. R. Webber, “QCD and collider physics,” *Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol.* **8** (1996) 1.

- [10] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert and J. Winter, “Event generation with SHERPA 1.1,” JHEP **0902** (2009) 007 [arXiv:0811.4622 [hep-ph]].
- [11] S. Schumann and F. Krauss, “A Parton shower algorithm based on Catani-Seymour dipole factorisation,” JHEP **0803** (2008) 038 [arXiv:0709.1027 [hep-ph]].
- [12] Y. L. Dokshitzer, G. D. Leder, S. Moretti and B. R. Webber, “Better jet clustering algorithms,” JHEP **9708**, 001 (1997) [hep-ph/9707323];
- [13] M. Wobisch and T. Wengler, “Hadronization corrections to jet cross sections in deep-inelastic scattering,” arXiv:hep-ph/9907280; M. Wobisch, “Measurement and QCD analysis of jet cross sections in deep-inelastic positron proton collisions at $s^{*}(1/2) = 300\text{-GeV}$,” DESY-THESIS-2000-049.
- [14] B. R. Webber, “QCD jets and parton showers,” arXiv:1009.5871 [hep-ph].
- [15] S. Catani, Y. L. Dokshitzer, F. Fiorani and B. R. Webber, “Average number of jets in e^+e^- annihilation,” Nucl. Phys. B **377** (1992) 445.
- [16] M. Cacciari, G. P. Salam and G. Soyez, “The Catchment Area of Jets,” JHEP **0804** (2008) 005 [arXiv:0802.1188 [hep-ph]].
- [17] E. Gerwick, T. Plehn, S. Schumann and P. Schichtel, “Scaling Patterns for QCD Jets,” JHEP **1210**, 162 (2012) [arXiv:1208.3676 [hep-ph]].
- [18] E. Gerwick, T. Plehn and S. Schumann, “Understanding Jet Scaling and Jet Vetos in Higgs Searches,” Phys. Rev. Lett. **108**, 032003 (2012) [arXiv:1108.3335 [hep-ph]].
- [19] C. Englert, T. Plehn, P. Schichtel and S. Schumann, “Establishing Jet Scaling Patterns with a Photon,” JHEP **1202** (2012) 030 [arXiv:1108.5473 [hep-ph]].
- [20] N. Brown and W. J. Stirling, “Jet cross-sections at leading double logarithm in e^+e^- annihilation,” Phys. Lett. B **252**, 657 (1990).
- [21] A. Banfi, G. Corcella and M. Dasgupta, “Angular ordering and parton showers for non-global QCD observables,” JHEP **0703**, 050 (2007).